# COVID-19 Dashboard – Comparing rates of infection across the populations of Victoria, Australia and Washington state, USA

## xxIntroduction to the problem

After successfully controlling the spread of COVID-19 from January to June 2020, Australian governments and residents faced a much more significant second wave of infections following a failure in quarantine procedure that was announced publicly on May 27. This second wave of infections has disproportionately affected the state of Victoria.

With cases now falling, our project aims to explore a number of hypotheses about the impact of demographic characteristics on the spread of COVID-19 in Victoria. Specifically:

* Do lower income areas have more cases?
* Are areas with higher aged populations more likely to have more cases?
* Are areas with higher education levels less affected? (Is there a difference?)
* Are people working in particular industries more likely to be infected?
* Is there a difference in the infection rate between the genders? What about the death rate?

Our approach is to design a dashboard that will include different views, each corresponding to different characteristics of the population (i.e. Income, Age, Gender Education, Employment, Occupation). See 'Sketch.jpg' file in our repository for an overview of the basic design. While our main case study is the state of Victoria, we will perform similar analyses on data depicting the COVID pandemic (in another similar state in the USA) across the USA, by way of comparison. These comparisons will allow us to explore further questions such as: Is the impact of sociodemographic factors such as age, gender, education, income and employment greater in the American setting than in Australia?

The COVID-19 cases data pack will be extracted from the official website of 'Department of Health and Human Services Victoria'. The census dataset will be generated from the ABS and ATO's database (see the links below or RESOURCES.md file in our repository for more details on the data sources). We will adopt the postcode of different districts in Victoria as the connection point for different databases. For the USA, we will source COVID case data from John Hopkins, and census data from the United States Census Bureau. Finally, we will use Bokeh, Geoplotlib or Dash to create data visualizations and present the results on an administrative map of Victoria.

## Scope and limitations of the study

This study was limited by the time allocated (4 weeks), and the scope of the course, which focused on data analysis and visualisation techniques using Python, but did not include any statistical analysis. As such, our findings are general in nature and have not undergone rigorous statistical interrogation.

### Limitations of the data

Large-scale social research, and particularly censuses, is costly and time-consuming to undertake. As a result, while population-level demographic data is publicly available, it is not regularly updated. In the USA, a population-wide census is conducted every ten years, with estimates projected for the intervening years. Unfortunately for this study, the next American census is open now, meaning the data used was based on projections that are now ten years old. In addition, survey participation is not comprehensive, meaning data was not available for all counties in Washington state across all demographic variables in the study; in particular, income and education level were available for 20 and xx counties, respectively.

Likewise, census data for the state of Victoria, Australia was limited by the interval of data collection; censuses are conducted every five years in Australia[[1]](#footnote-1).

## Findings

### Age

Xx

### Gender

Xx

### Income

#### Data sources:

Median weekly household income from Australian Tax Office? xx check files/notes in repo “2016 VIC median\_age & median\_income.csv”

<https://www.dhhs.vic.gov.au/victorian-coronavirus-covid-19-data>

“Income and xx”, US Census Bureau 2019

In Victoria, median household income in each postcode did not seem to have an effect on the rate of infection with COVID-19. In Washington state, infection rates were much higher, with a mean infection rate across all counties of xx (mean infection rate in Victoria = xx). There appears to be a skew towards counties with lower median incomes, however the data is limited, as described above under *Limitations of the data*.

### Education

Xx

### Occupation

Xx

## Future directions

xx

1. https://www.abs.gov.au/websitedbs/censushome.nsf/home/about. [↑](#footnote-ref-1)